A Comparative Study of Some Modifications of CG Methods Under Exact Line Search

Yasir Salih[a]*, Mustafa Mamat[b], and Sukono[c]*

[a]*Department of Mathematics, Faculty of Education, Red Sea University, SUDAN.*
[a,b]*Faculty of Informatics and Computing, Universiti Sultan ZainalAbidin, Terengganu, MALAYSIA.*
[c]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Bandung, INDONESIA*

E-mail address: yasirsalih2015@gmail.com; sukono@unpad.ac.id

Abstract

Conjugate Gradient (CG) method is a technique used in solving nonlinear unconstrained optimization problems. In this paper, we analysed the performance of two modifications β_k^{YHM} and β_k^{MRM} and compared the results with the classical conjugate gradient methods of β_k^{FR} and β_k^{PRP}. These proposed methods possesse global convergence properties for general functions using exact line search. Numerical experiments show that the two modifications are more efficient for the test problems compared to classical CG coefficients.

Keywords: Conjugate Gradient Method, unconstrained optimization, exact line search

1. Introduction

Consider the following nonlinear unconstrained optimization problems
$$\min\{f(x) : x \epsilon R^n\}, \tag{1}$$

where $f : R^n \to R$ is a continuous differentiable function that is bounded below. To solve (1), starting from an initial point $x_0 \epsilon R^n$, we obtained the next iterative point as follows
$$x_{k+1} = x_k + \alpha_k d_k \ , \ \ k = 0,1,2, \dots \tag{2}$$

where α_k is the step-length and d_k is the search direction defined as;
$$d_k = \begin{cases} -g_k, & if k = 0 \\ -g_k + \beta_k d_{k-1}, & if k \geq 1 \end{cases} \tag{3}$$

The scalar β_k is known as the conjugate gradient parameter and the gradient $g_k = \nabla f(x_k)$. Some of the classical CG parameters are Fletcher and Reeves (FR) (Fletcher and Reeves,1964) Polak-Ribiere-Polyak (PRP) (Polak and Ribiere, 1969; Polyak, 1969) Hestenes-Steifel (HS) (Hestenes and Stiefel,1952) Conjugate Descent (CD) (Fletcher, 1980), Liu-Storey (LS) method (Liu and Storey, 1991). These parameters and are defined as

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \qquad \beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\beta_{k-1}^T g_{k-1}}$$

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_k^T(g_k - g_{k-1})}, \qquad \beta_k^{CD} = \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}},$$

$$\beta_k^{LS} = \frac{g_k^T(g_k - g_{k-1})}{-d_{k-1}^T g_{k-1}}, \qquad \beta_k^{DY} = \frac{\|g_k^2\|}{d_{k-1}^T(g_k - g_{k-1})}$$

Most of these classical methods may not generally be convergent but often possess good convergence property. Recently, more research has been published on new modifications of CG methods. For good references of recent CG methods with significant results, refer to Sulaiman et al. (Sulaiman el al, 2018; Sulaiman et al, 2015; Kamilu et al, 2018; Yasir et al, 2018). Thus, in this paper, we analyszed the performance of two modifications of CG coefficients and compared the perfomance with that of classical CG methods of FR and PRP under exact line search. This is done to improve the overall performance of the resulting algorithms.

2. Two Modifications and Algorithm

In this section, we present two modifications known as β_k^{YHM} (Saliha et al, 2018) and β_k^{MRM} (Hamoda et al, 2017) and defined by,

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} & if \ \ 0 \le g_k^T g_{k-1} \le \|g_k\|^2 \\[3mm] \dfrac{g_k^T\left(g_k - \dfrac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2} & otherwise \end{cases} \qquad (4)$$

$$\beta_k^{MRM} = \frac{g_k^T\left(g_k - \dfrac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \qquad (5)$$

The following algorithm is a general algorithm for solving unconstrained optimization problems.
Algorithm
Step1: Given $x_0 \epsilon R^n$, set $\varepsilon > 0$ $d_0 = -g_0$, If $\|g_k\| = 0$, then stop. Else,
Step2: Computing step size by,
$$\alpha_k = arg\ min_{\alpha \ge 0} f(x_k + \alpha d_k),$$
Step3: Compute
$$x_{k+1} = x_k + \alpha_k d_k.$$

Step4: Compute β_k and generate d_{k+1} by (3)

Step5: Convergence and stopping criteria. Check if$\|g_{k+1}\| \leq \epsilon$stop. Else, Set $k = k + 1$go to Step 2

3. Global Convergence of the New Modifications

In this section, we prove the global convergence of$\boldsymbol{\beta}_k^{YHM}$and $\boldsymbol{\beta}_k^{MRM}$under exact line search. We begin with the sufficient descent condition.

3.1 Sufficient Descent Condition

For sufficient descent condition to hold,

$$g_k^T d_k \leq -c\|g_k\|^2 \ \ \forall k \geq 0 \ and \ c > 0. \tag{6}$$

The following theorem is used to show our new modification possess with sufficient descent condition under exact line search.

Theorem 1. Let $\{x_k\}$ and $\{d_k\}$be sequence generated from (2), (3) and the above algorithm, where the step size α_k is determined by the exact line search. Then (6) holds for all k $\geq$ 0.

Proof. The prove of this theorem can be found in (Saliha et al, 2018; Hamoda et al, 2017).

3.2 Global Convergence Properties

In this section, we prove the global convergence properties of the new modification under some assumptions.

Assumption 1

(I) $f(x)$ is bounded from below on the level set R^n and is continuous and differentiable in a neighbourhood N of the set $N = \{x \in R^n \ , \ f(x) \leq f(x_0)\}$ at the initial point x$_0$.

(II) The gradient $g(x)$ is Lipschitz continuous in N, so there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \leq L\|x - y\|, \ \forall x, y \in N.$

The following Lemma by Zoutendijk (Zoutendijk, 1970) is used to prove the global convergence.

Lemma 1.

Suppose that Assumption 1 holds true. Consider any CG method of the form (3) where d_k is the search direction. Then, Zoutendijk condition holds, that is,

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{7}$$

The following theorem is based on Lemma 1.

Theorem 2

Suppose assumption 1 holds true, consider β_k^{YHM} and β_k^{MRM} methods of the form (2) and (3) where$\alpha_k$is obtained using exact line search, then,

$$\lim_{k\to\infty} \inf \|g_k\| = 0 \tag{8}$$

Proof. The prove of this theorem can be found in (Saliha et al, 2018; Hamoda et al, 2017).

4. Numerical Results

In this section, we report the detailed numerical results based on the comparisons of our new modifications with classical CG Algorithms of FR and PRP. All algorithms are implemented under exact line search. We selected 31 test functions from Andrei (Andrei, 2008) with different dimensions. We considered $\varepsilon = 10^{-6}$ and $\|g_k\| \leq 10^{-6}$ to be the stopping criteria as suggested by Hillstrom (Hillstrom, 1977). In all cases, we used four initial points. These four initial points lead us to test the global convergence of the new modifications. All algorithms are coded on MATLAB Version R2014a. The test was run on an Intel(R) Core™ i5-M520 (2.40GHz), 4GB for RAM memory and Windows 7 Professional operating system. The Numerical results are based on number of iterations and CPU time as presented in Table 1. Also, the performance results of these methods are shown in Figure 1 and Figure 2 respectively, using a performance profile introduced by Dolan and More (Dolan and Moré, 2002).

Table 1: A list of problem functions.

No	Function	Dimension	Initial points
1	Six hump camel	2	-10, -8, 8, 10
2	Booth	2	10, 25, 50, 100
3	Treccani	2	5, 10, 20, 50
4	Zettl	2	5, 10, 20, 30
5	Ex –rosenbrock	2,4,10,100,500,1000,10000	13, 25, 30, 50
6	Extended penalty	2,4,10,100	50, 60, 70, 80
7	Generalized tridiagonal 1	2,4,10,100	30, 35, 40, 45
8	Shalow	2,4,10,100,500,1000,10000	10, 25, 50, 70
9	Ex - tridiagonal1	2,4,10,100,500,1000,10000	12, 17, 20, 30
10	Extended white and holst	2,4,10,100,500,1000,10000	3, 10, 30, 50
11	Quadrtic QF2	2,4,10,100,500,1000	10, 30, 50, 100
12	Extended denschnb	2,4,10,100,500,1000,10000	8, 13, 30, 50

| 13 | Hager | 2,4,10,100 | 1, 3, 5, 7 |
14	Ex-powell	4,8,20,100,500,1000	-1, 1, 7, 11
15	Extended beale	2,4,10,100,500,1000,10000	-19, 1, 13, 23
16	Ex –himmelblau	100,500,1000,10000	50, 70, 100, 125
17	Diagonal 2	2,4,10,100,500,1000	-1, 1, 2, 3
18	Perturbed quadratic	2,4,10,100,500,1000	1, 5, 10, 15
19	Sum Squares function	2,4,10,100,500,1000	1, 10, 20, 30
20	Ex- quadratic penalty QP2	4,10,100,500,1000,10000	17, 18, 19, 20
21	Raydan1 function	2,4,10,100	1, 3, 5, 7
22	Generalized tridiagonal 2	2,4,10,100	1, 7, 8, 14
23	Quadratic QF1	2,4,10,100,500,1000	1, 2, 3, 4
24	Dixon and Price	2,4,10,100	100, 125, 150, 175
25	Fletcher	4,10,100,500,1000	7, 11, 13, 15
26	Ex- maratos	2,4,10,100	5, 10, 12, 15
27	Leon function	2	2, 5, 8, 10
28	Extended wood	4	3, 5, 20, 30
29	Quartic function	4	5, 10, 15, 20
30	Matyas function	2	5, 10, 15, 20
31	Colville function	4	2, 4, 7, 10
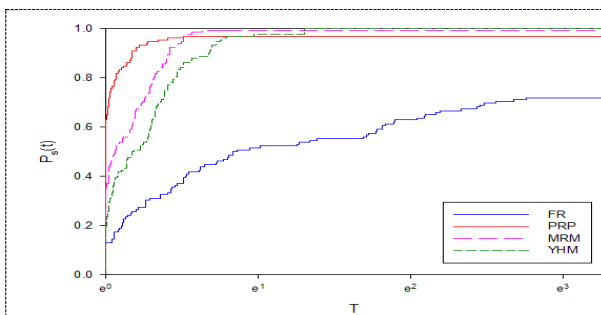
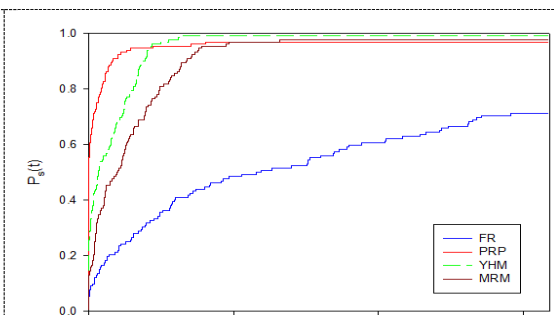| **Figure 1:** Performance profile based on number of iteration. | **Figure 2.** Performance profile based on CPU Time. |

From the above figures, it can be seen that the new modifications are better than the classical methods of FR and PRP based on number of iteration and CPU times. The results show that the proposed method solves 99.4% and 98% of the test problems respectively. Meanwhile, FR method solves about 72% while PRP solves 97% of the test problems. Hence, we can say that β_k^{YHM} and β_k^{MRM} are more efficient with robust performance.

5. Conclusion

In this paper, we presented the performance of two new modifications for nonlinear unconstrained optimization and compared them with the classical CG coefficient of FR and PRP methods under exact line search. These proposed methods possess the global convergence condition under some assumptions and satisfy the sufficient descent condition. Numerical results showed that the proposed methods outperformed the classical methods of FR and PRP in solving the standard unconstrained optimization problems.

Acknowledgements

References

Andrei N. (2008). An unconstrained optimization test functions collection. *Adv. Model. Optim*, 10(1): p. 147-161.

Dolan E.D and Moré JJ. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming*, 91(2): p. 201-213.

Fletcher R and Reeves C M. (1964). Function minimization by conjugate gradients. *The computer journal*, 7(2): p. 149-154.

Fletcher R. 1980. Practical methods of optimization. Vol. 1, Unconstrained Optimization.

Hamoda M, Rivaie M and Mamat M. (2017). A New Nonlinear Conjugate Gradient Method With Exact Line Search For Unconstrained Optimization. *Journal of Humanities and Applied Science (JHAS)*, pp. 1-16.

Hestenes MR and Stiefel E. (1952). Methods of conjugate gradients for solving linear systems. Vol. 49: NBS.

Hillstrom K.E. (1977). A simulation test approach to the evaluation of nonlinear optimization algorithms. *ACM Transactions on Mathematical Software (TOMS),* 3(4): p. 305-315. 1977. 3(4): p. 305-315.

Kamilu, K., Sulaiman, I. M., Waziri, M. Y., and Abashar, A. (2018). Another Improved Three Term PRP-CG Method with Global Convergent Properties for Solving Unconstrained Optimization Problems. *Malaysian Journal of Computing and Applied Mathematics*. 1(1), 1-10.

Liu Y and Storey C. 1991.Efficient generalized conjugate gradient algorithms, part 1: theory. *Journal of optimization theory and applications*, 69(1): p. 129-137.

Polyak B.T. (1969). The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, 9(4): p. 94-112

Polak E and Ribiere G. (1969). Note sur la convergence de méthodes de directions conjuguées. Revue française d'informatique et de recherche opérationnelle. *Série rouge*, 3(16): p. 35-43.

Saliha, Y., M.A. Hamoda, and M. Rivaiec. (2018). New Hybrid Conjugate Gradient Method with Global Convergence Properties for Unconstrained Optimization. *Malaysian Journal of Computing and Applied Mathematics*, 1(1): p. 29-38

Sulaiman, I. M., Mamat, M., Abashar, A., and Zabidin, S. (2015). A Modified Nonlinear Conjugate Gradient Method for Unconstrained Optimization, Applied Mathematical Sciences 9(54), 2671 – 2682.

Sulaiman, I. M., Waziri, M. Y., Olowo, E. S., and Talat, A. N. (2018). Solving Fuzzy Nonlinear Equations with a New Class of Conjugate Gradient Method. Malaysian Journal of Computing and Applied Mathematics. 1(1), 11-19.

Yasir, S., Mohamed, A. H., and Rivaie, M. (2018). New Hybrid Conjugate Gradient Method with Global Convergent Properties for Solving Unconstrained Optimization. Malaysian Journal of Computing and Applied Mathematics. 1(1), 29-38.

Zoutendijk G. (1970). Nonlinear programming, computational methods. *Integer and nonlinear programming*, 143(1): p. 37-8